



OIPE

## RAW SEQUENCE LISTING

DATE: 02/04/2002

PATENT APPLICATION: US/09/978,242

TIME: 11:38:59

Input Set : N:\CrF3\RULE60\09978242.raw

Output Set: N:\CRF3\02042002\I978242.raw

## SEQUENCE LISTING

## 3 (1) GENERAL INFORMATION:

5 (i) APPLICANT: Bandman, Olga  
 6 Yue, Henry  
 7 Corley, Neil C.  
 8 Shah, Purvi

10 (ii) TITLE OF INVENTION: HUMAN NUCLEOLIN-LIKE PROTEIN

12 (iii) NUMBER OF SEQUENCES: 3

14 (iv) CORRESPONDENCE ADDRESS:

15 (A) ADDRESSEE: Incyte Pharmaceuticals, Inc.  
 16 (B) STREET: 3174 Porter Drive  
 17 (C) CITY: Palo Alto  
 18 (D) STATE: CA  
 19 (E) COUNTRY: USA  
 20 (F) ZIP: 94304

22 (v) COMPUTER READABLE FORM:

23 (A) MEDIUM TYPE: Diskette  
 24 (B) COMPUTER: IBM Compatible  
 25 (C) OPERATING SYSTEM: DOS  
 26 (D) SOFTWARE: FastSEQ for Windows Version 2.0

28 (vi) CURRENT APPLICATION DATA:

C--&gt; 29 (A) APPLICATION NUMBER: US/09/978,242

C--&gt; 30 (B) FILING DATE: 15-Oct-2001

31 (C) CLASSIFICATION:

33 (vii) PRIOR APPLICATION DATA:

35 (A) APPLICATION NUMBER: 09/241,333

36 (B) FILING DATE:

38 (A) APPLICATION NUMBER: 08/990,114

39 (B) FILING DATE:

41 (viii) ATTORNEY/AGENT INFORMATION:

42 (A) NAME: Billings, Lucy J.

43 (B) REGISTRATION NUMBER: 36,749

44 (C) REFERENCE/DOCKET NUMBER: PF-0451 US

46 (ix) TELECOMMUNICATION INFORMATION:

47 (A) TELEPHONE: 650-855-0555

48 (B) TELEFAX: 650-845-4166

49 (C) TELEX:

52 (2) INFORMATION FOR SEQ ID NO: 1:

54 (i) SEQUENCE CHARACTERISTICS:

55 (A) LENGTH: 545 amino acids

56 (B) TYPE: amino acid

57 (C) STRANDEDNESS: single

58 (D) TOPOLOGY: linear

ENTERED

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/978,242

DATE: 02/04/2002

TIME: 11:38:59

Input Set : N:\Crf3\RULE60\09978242.raw

Output Set: N:\CRF3\02042002\I978242.raw

```

60      (vii) IMMEDIATE SOURCE:
61          (A) LIBRARY: TLYMNOT05
62          (B) CLONE: 2809795
64      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
66 Met Ala Thr Glu His Val Asn Gly Asn Gly Thr Glu Glu Pro Met Asp
67   1          5          10          15
68 Thr Thr Ser Ala Val Ile His Ser Glu Asn Phe Gln Thr Leu Leu Asp
69          20          25          30
70 Ala Gly Leu Pro Gln Lys Val Ala Glu Lys Leu Asp Glu Ile Tyr Val
71          35          40          45
72 Ala Gly Leu Val Ala His Ser Asp Leu Asp Glu Arg Ala Ile Glu Ala
73          50          55          60
74 Leu Lys Glu Phe Asn Glu Asp Gly Ala Leu Ala Val Leu Gln Gln Phe
75          65          70          75          80
76 Lys Asp Ser Asp Leu Ser His Val Gln Asn Lys Ser Ala Phe Leu Cys
77          85          90          95
78 Gly Val Met Lys Thr Tyr Arg Gln Arg Glu Lys Gln Gly Thr Lys Val
79          100         105         110
80 Ala Asp Ser Ser Lys Gly Pro Asp Glu Ala Lys Ile Lys Ala Leu Leu
81          115         120         125
82 Glu Arg Thr Gly Tyr Thr Leu Asp Val Thr Thr Gly Gln Arg Lys Tyr
83          130         135         140
84 Gly Gly Pro Pro Pro Asp Ser Val Tyr Ser Gly Gln Gln Pro Ser Val
85          145         150         155         160
86 Gly Thr Glu Ile Phe Val Gly Lys Ile Pro Arg Asp Leu Phe Glu Asp
87          165         170         175
88 Glu Leu Val Pro Leu Phe Glu Lys Ala Gly Pro Ile Trp Asp Leu Arg
89          180         185         190
90 Leu Met Met Asp Pro Leu Thr Gly Leu Asn Arg Gly Tyr Ala Phe Val
91          195         200         205
92 Thr Phe Cys Thr Lys Glu Ala Ala Gln Glu Ala Val Lys Leu Tyr Asn
93          210         215         220
94 Asn His Glu Ile Arg Ser Gly Lys His Ile Gly Val Cys Ile Ser Val
95          225         230         235         240
96 Ala Asn Asn Arg Leu Phe Val Gly Ser Ile Pro Lys Ser Lys Thr Lys
97          245         250         255
98 Glu Gln Ile Leu Glu Glu Phe Ser Lys Val Thr Glu Gly Leu Thr Asp
99          260         265         270
100 Val Ile Leu Tyr His Gln Pro Asp Asp Lys Lys Lys Asn Arg Gly Phe
101          275         280         285
102 Cys Phe Leu Glu Tyr Glu Asp His Lys Thr Ala Ala Gln Ala Arg Arg
103          290         295         300
104 Arg Leu Met Ser Gly Lys Val Lys Val Trp Gly Asn Val Gly Thr Val
105          305         310         315         320
106 Glu Trp Ala Asp Pro Ile Glu Asp Pro Asp Pro Glu Val Met Ala Lys
107          325         330         335
108 Val Lys Val Leu Phe Val Arg Asn Leu Ala Asn Thr Val Thr Glu Glu
109          340         345         350
110 Ile Leu Glu Lys Ala Phe Ser Gln Phe Gly Lys Leu Glu Arg Val Lys

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/978,242

DATE: 02/04/2002

TIME: 11:38:59

Input Set : N:\Crif3\RULE60\09978242.raw

Output Set: N:\CRF3\02042002\I978242.raw

```

111          355          360          365
112 Lys Leu Lys Asp Tyr Ala Phe Ile His Phe Asp Glu Arg Asp Gly Ala
113          370          375          380
114 Val Lys Ala Met Glu Glu Met Asn Gly Lys Asp Leu Glu Gly Glu Asn
115          385          390          395          400
116 Ile Glu Ile Val Phe Ala Lys Pro Pro Asp Gln Lys Arg Lys Glu Arg
117          405          410          415
118 Lys Ala Gln Arg Gln Ala Ala Lys Asn Gln Met Tyr Asp Asp Tyr Tyr
119          420          425          430
120 Tyr Tyr Gly Pro Pro His Met Pro Pro Pro Thr Arg Gly Arg Gly Arg
121          435          440          445
122 Gly Gly Arg Gly Gly Tyr Gly Tyr Pro Pro Asp Tyr Tyr Gly Tyr Glu
123          450          455          460
124 Asp Tyr Tyr Asp Tyr Tyr Gly Tyr Asp Tyr His Asn Tyr Arg Gly Gly
125          465          470          475          480
126 Tyr Glu Asp Pro Tyr Tyr Gly Tyr Glu Asp Phe Gln Val Gly Ala Arg
127          485          490          495
128 Gly Arg Gly Gly Arg Gly Ala Arg Gly Ala Ala Pro Ser Arg Gly Arg
129          500          505          510
130 Gly Ala Ala Pro Pro Arg Gly Arg Ala Gly Tyr Ser Gln Arg Gly Gly
131          515          520          525
132 Pro Gly Ser Ala Arg Gly Val Arg Ala Gly Lys Arg Gly Arg Gly Arg
133          530          535          540
134 Ser
135 545

```

137 (2) INFORMATION FOR SEQ ID NO: 2:

139 (i) SEQUENCE CHARACTERISTICS:

140 (A) LENGTH: 2079 base pairs

141 (B) TYPE: nucleic acid

142 (C) STRANDEDNESS: single

143 (D) TOPOLOGY: linear

145 (vii) IMMEDIATE SOURCE:

146 (A) LIBRARY: TLYMNOT05

147 (B) CLONE: 2809795

149 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

```

151 GGGCGCGCGC GCGCACCGGG AGCGCGCTCG GAGGCGAGTG GAACTGGATC GGGTTTGCTG      60
152 CCAGCGGCGT GAGCTTCGGC CGGCATTTTA CAACAGCTCC ACTCGCGCCG GACACAGGGA      120
153 GCAGCGAGCA CGCGTTTCCC GCAACCCGAT ACCATCGGAC AGGATTCTC CGCCTCAGCC      180
154 CAACGGGGGAG ATCTCTGGAA ACATGGCTAC AGAACATGTT AATGGAAATG GTACTGAAGA      240
155 GCCCATGGAT ACTACTTCTG CAGTTATCCA TTCAGAAAAT TTTCAGACAT TGCTTAGATGC      300
156 TGGTTTACCA CAGAAAGTTG CTGAAAAACT AGATGAAATT TACGTTGCAG GGCTAGTTGC      360
157 ACATAGTGAT TTAGATGAAA GAGCTATTGA AGCTTTAAAA GAATTCAATG AAGACGGTGC      420
158 ATTGGCAGTT CTTCACAGT TTAAAGACAG TGATCTCTCT CATGTTCAGA ACAAAGTGTC      480
159 CTTTTTATGT GGAGTCATGA AGACTTACAG GCAGAGAGAA AAACAAGGGA CCAAAGTAGC      540
160 AGATTCTAGT AAAGGACCAG ATGAGGCAAA AATTAAGGCA CTCTTGGAAG GAACAGGCTA      600
161 CACACTTGAT GTGACCACTG GACAGAGGAA GTATGGAGGA CCACCTCCAG ATTCCGTTTA      660
162 TTCAGGTCAG CAGCCTTCTG TTGGCACTGA GATATTTGTG GGAAAGATCC CAAGAGATCT      720
163 ATTTGAGGAT GAACTTGTTT CATTATTTGA GAAAGCTGGA CCTATATGGG ATCTTCGTCT      780
164 AATGATGGAT CCACTCACTG GTCTCAATAG AGGTTATGCG TTTGTCACCT TTTGTACAAA      840

```

## RAW SEQUENCE LISTING

DATE: 02/04/2002

PATENT APPLICATION: US/09/978,242

TIME: 11:38:59

Input Set : N:\Crif3\RULE60\09978242.raw

Output Set: N:\CRF3\02042002\I978242.raw

```

165 AGAAGCAGCT CAGGAGGCTG TTAAACTGTA TAATAATCAT GAAATTCGTT CTGGAAAACA 900
166 TATTGGTGTC TGCAATCAG TTGCCAACAA TAGGCTTTTT GTGGGCTCTA TTCCTAAGAG 960
167 TAAAACCAAG GAACAGATTC TTGAAGAATT TAGCAAAGTA ACAGAGGGTC TTACAGACGT 1020
168 CATTTTATAC CACCAACCGG ATGACAAGAA AAAAAACAGA GGCTTTTGCT TTCTTGAATA 1080
169 TGAAGATCAC AAAACAGCTG CCCAGGCAAG GCGTAGGTTA ATGAGTGGTA AAGTCAAGGT 1140
170 CTGGGGGAAT GTTGGAACTG TTGAATGGGC TGATCCTATA GAAGATCCTG ATCCTGAGGT 1200
171 TATGGCAAAG GTAAAAGTGC TGTTTGACG CAACCTTGCC AATACTGTAA CAGAAGAGAT 1260
172 TTTAGAAAAG GCATTTAGTC AGTTTGGGAA ACTGGAACGA GTGAAGAAGT TAAAAGATTA 1320
173 TGC GTTCATT CATTTTGATG AGCGAGATGG TGCTGTCAAG GCTATGGAAG AAATGAATGG 1380
174 CAAAGACTTG GAGGGAGAAA ATATTGAAAT TGTTTTTGCC AAGCCACCAG ATCAAGAAAAG 1440
175 GAAAGAAAGA AAAGCTCAGA GGCAAGCAGC AAAAAATCAA ATGTATGACG ATTACTACTA 1500
176 TTATGGTCCA CCTCATATGC CCCCTCCAAC AAGAGGTCGA GGGCGTGGAG GTAGAGGTGG 1560
177 TTATGGATAT CCTCCAGATT ATTATGGATA TGAAGATTAT TATGATTATT ATGTTTATGA 1620
178 TTACCATAAC TATCGTGGTG GATATGAAGA TCCATACTAT GGTATGAAG ATTTTCAAGT 1680
179 TGGAGCTAGA GGAAGGGGTG GTAGAGGAGC AAGGGGTGCT GCTCCATCCA GAGGTCGTGG 1740
180 GGCTGCTCCT CCCC GCGGTA GAGCCGGTTA TTCACAGAGA GGAGGTCCTG GATCAGCAAG 1800
181 AGGCGTTCGA GCAGGGAAAA GGGGTGAGG CCGGTCCTGA CCTGTTACAA TGAAGACTGA 1860
182 CTTGCTATGT GGGATTACAC CAGAAGCTTG CAGTGGAGTA ATGGTAAGGA AATCAAGCAA 1920
183 CCTTAAATAT GTCGGCTGTA TAGGAGCATA TTCTATTGCA GAAGACCTTC CTATGAAGAT 1980
184 CATGGAATCA AATACGGGAC ATTGAATAA TACTTGGACT TTGATATGAA TTTCTTTAAC 2040
185 AATTTTCTCT GCAGTGCAAG TTATTAACCT AAAGCTACT 2079

```

187 (2) INFORMATION FOR SEQ ID NO: 3:

189 (i) SEQUENCE CHARACTERISTICS:

190 (A) LENGTH: 714 amino acids

191 (B) TYPE: amino acid

192 (C) STRANDEDNESS: single

193 (D) TOPOLOGY: linear

195 (vii) IMMEDIATE SOURCE:

196 (A) LIBRARY: GenBank

197 (B) CLONE: 128842

199 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

```

201 Met Val Lys Leu Ala Lys Ala Gly Lys Thr His Gly Glu Ala Lys Lys
202 1 5 10 15
203 Met Ala Pro Pro Lys Glu Val Glu Asp Ser Glu Asp Glu Glu
204 20 25 30
205 Met Ser Glu Glu Glu Asp Asp Ser Ser Gly Glu Glu Val Val Ile Pro
206 35 40 45
207 Gln Lys Lys Gly Lys Lys Ala Thr Ala Thr Pro Ala Lys Lys Val Val
208 50 55 60
209 Val Ser Gln Thr Lys Lys Val Ala Val Pro Thr Pro Ala Lys Lys Ala
210 65 70 75 80
211 Ala Val Thr Pro Gly Lys Lys Ala Ala Ala Thr Pro Ala Lys Lys Ala
212 85 90 95
213 Val Thr Pro Ala Lys Ala Val Ala Thr Pro Gly Lys Lys Gly Ala Thr
214 100 105 110
215 Gln Ala Lys Ala Leu Val Ala Thr Pro Gly Lys Lys Gly Ala Val Thr
216 115 120 125
217 Pro Ala Lys Gly Ala Lys Asn Gly Lys Asn Ala Lys Lys Glu Asp Ser
218 130 135 140

```

## RAW SEQUENCE LISTING

DATE: 02/04/2002

PATENT APPLICATION: US/09/978,242

TIME: 11:38:59

Input Set : N:\Crf3\RULE60\09978242.raw

Output Set: N:\CRF3\02042002\I978242.raw

```

219 Asp Glu Asp Glu Asp Asp Asp Asp Asp Glu Asp Asp Ser Asp Glu Asp
220 145 150 155 160
221 Glu Glu Asp Glu Glu Glu Asp Glu Phe Glu Pro Pro Val Val Lys Gly
222 165 170 175
223 Lys Gln Gly Lys Val Ala Ala Ala Ala Pro Ala Ser Glu Asp Glu Asp
224 180 185 190
225 Glu Glu Glu Asp Glu Glu Glu Glu Glu Glu Asp Glu Glu Glu Glu Asp
226 195 200 205
227 Asp Ser Glu Glu Glu Glu Ala Met Glu Ile Thr Pro Ala Lys Gly Lys
228 210 215 220
229 Lys Ala Pro Ala Lys Val Val Pro Val Lys Ala Lys Asn Val Ala Glu
230 225 230 235 240
231 Glu Asp Asp Asp Asp Glu Glu Glu Asp Glu Asp Glu Glu Glu Asp Glu
232 245 250 255
233 Glu Glu Glu Glu Asp Glu Glu Glu Glu Glu Glu Glu Glu Glu Glu
234 260 265 270
235 Pro Val Lys Pro Ala Pro Gly Lys Arg Lys Lys Glu Met Thr Lys Gln
236 275 280 285
237 Lys Glu Val Pro Glu Ala Lys Lys Gln Lys Val Glu Gly Ser Glu Ser
238 290 295 300
239 Thr Thr Pro Phe Asn Leu Phe Ile Gly Asn Leu Asn Pro Asn Lys Ser
240 305 310 315 320
241 Val Ala Glu Leu Lys Val Ala Ile Ser Glu Pro Phe Ala Lys Asn Asp
242 325 330 335
243 Leu Ala Val Val Asp Val Arg Thr Gly Thr Asn Arg Lys Phe Gly Tyr
244 340 345 350
245 Val Asp Phe Glu Ser Ala Glu Asp Leu Glu Lys Ala Leu Glu Leu Thr
246 355 360 365
247 Gly Leu Lys Val Phe Gly Asn Glu Ile Lys Leu Glu Lys Pro Lys Gly
248 370 375 380
249 Arg Asp Ser Lys Lys Val Arg Ala Ala Arg Thr Leu Leu Ala Lys Asn
250 385 390 395 400
251 Leu Ser Phe Asn Ile Thr Glu Asp Glu Leu Lys Glu Val Phe Glu Asp
252 405 410 415
253 Ala Leu Glu Ile Arg Leu Val Ser Gln Asp Gly Lys Ser Lys Gly Ile
254 420 425 430
255 Ala Tyr Ile Glu Phe Lys Ser Glu Ala Asp Ala Glu Lys Asn Leu Glu
256 435 440 445
257 Glu Lys Gln Gly Ala Glu Ile Asp Gly Arg Ser Val Ser Leu Tyr Tyr
258 450 455 460
259 Thr Gly Glu Lys Gly Gln Arg Gln Glu Arg Thr Gly Lys Asn Ser Thr
260 465 470 475 480
261 Trp Ser Gly Glu Ser Lys Thr Leu Val Leu Ser Asn Leu Ser Tyr Ser
262 485 490 495
263 Ala Thr Glu Glu Thr Leu Gln Glu Val Phe Glu Lys Ala Thr Phe Ile
264 500 505 510
265 Lys Val Pro Gln Asn Gln Gln Gly Lys Ser Lys Gly Tyr Ala Phe Ile
266 515 520 525
267 Glu Phe Ala Ser Phe Glu Asp Ala Lys Glu Ala Leu Asn Ser Cys Asn

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/978,242

DATE: 02/04/2002

TIME: 11:39:00

Input Set : N:\Crf3\RULE60\09978242.raw

Output Set: N:\CRF3\02042002\I978242.raw

L:29 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]

L:30 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]